

空间数据质量模型研究

杜道生

(武汉测绘科技大学测绘遥感信息工程国家重点实验室, 武汉 430079)

王占宏 马聪丽

(国家测绘局测绘标准化研究所, 西安 710054)

摘要 空间数据对于地理信息系统就如同汽油对于汽车一样重要. 空间数据的质量更是地理信息系统分析结果是否准确和合理的重要因素. 因此根据质量评价方法的不同, 提出了相应的空间数据质量模型, 并进而提出空间数据质量的度量模型, 这对科学、合理地进行数据质量评定及准确、统一的质量表达将起到概念级指导意义.

关键词 空间数据 质量度量 质量评价

中图法分类号: P208 N3 文献标识码: A 文章编号: 1006-8961(2000)07-0559-04

Study on Spatial Data Quality Model

DU Dao-sheng

(State Key Lab. for Information Engineering in Surveying and Mapping, WTUSM, Wuhan 430079)

WANG Zhan-hong, MA Cong-li

(Institute of Standardization in Surveying and Mapping, Xi'an 710054)

Abstract Spatial data for geographic information system is as important as the petrol for vehicle. Spatial data quality is more important factor for evaluating the analysis results if they are correct or reasonable in GIS. In this paper, a spatial data quality model is presented according to the method used for evaluating the spatial data quality, and a measurement model of spatial data quality is also given. Both two models, which are used for scientific and reasonable assessment of data quality as well as exact and unitary representation of data quality, are of great importance at conceptual level.

Keywords Spatial data, Quality measure, Quality evaluation

0 引言

数据是构成地理信息系统的四大要素(硬件、软件、数据、人员)之一. 地理信息系统中的数据包括空间数据和属性数据, 其中, 空间数据是地理信息系统的基础, 是贯穿于地理信息系统各个开发阶段、各个组成部分的极其重要的要素, 因此, 空间数据质量的优劣直接影响到地理信息系统各种分析结果的准确性与正确性, 进而关系到地理信息系统建立的成败.

目前, 对空间数据质量的科学度量及精确表达, 已成为数据提供者、使用者都非常关心的问题. 近十几年来, 国内外科学工作者对空间数据质量的研究高度重视, 并开展了深入的研究, 虽已提出了一些度量指标和数学模型, 但主要侧重于空间数据位置精度方面^[1~6]. 美国数字制图标准国家委员会(NCDCDS)、国际制图协会(ICA)空间数据质量委员会、国际标准化组织地理信息标准技术委员会(ISO/TC211)等组织均先后提出了一些数据质量元素, 但目前尚无统一的数据质量模型^[7~10].

实际上,质量评价方法不同,就决定了其采用不同的质量模型,但不论质量模型评价中采用什么评价方法,质量的最终表达应该使用统一的质量元素. 本文通过对质量评价方法的描述,提出了不同的空间数据质量模型,并进而提出了统一的空间数据质量的度量模型. 这对科学、合理地进行数据质量评定及准确、统一的质量表达,以及空间数据质量标准的制定将起到概念级指导意义.

1 空间数据质量的评价方法

一般,空间数据质量评价方法分直接评价方法和间接评价方法. 其中,直接评价方法是通过全面检测或抽样检测方式对数据集(产品)进行评价的方法,又称验收度量;间接评价方法是对数据的来源和质量、生产方法等间接信息进行数据集质量评价的方法,又称预估度量,它们之间的关系如图1所示. 这两种方法本质区别是面向的对象不同,其中直接评价方法面对的是生产出的数据集;而间接质量评价方法面对的是一些间接信息,因此只能通过误差传播的原理,根据间接信息估算出最终成品数据集的质量.

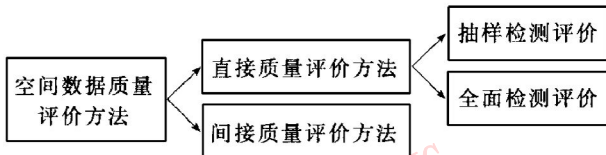


图1 空间数据质量评价方法

由此可见,在对空间数据质量进行评价时,首先要选择评价方法,然后根据评价方法和质量规范再确定质量特性(元素),并对数据集进行质量评价,最后形成质量评价报告,其过程如图2所示. 在实际操作时,数据质量评价方法的优选顺序为:先直接评价方法,后间接评价方法;或以直接评价方法为主,间接评价方法为辅.

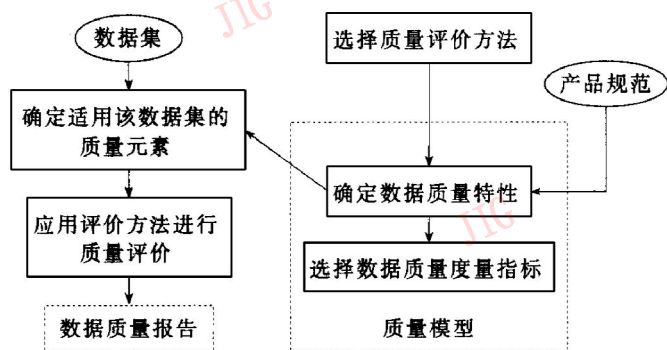


图2 空间数据质量评价过程

2 空间数据质量模型

空间数据分为矢量数据和栅格数据,其中,栅格数据又分为影像数据和格网数据. 不同类型的数据质量应采用不同的数据质量模型来评价. 由于空间数据质量是众多影响因素共同作用的结果,因此质量模型的确立,应在充分考虑这些因素后再确立其质量特性. 本文提出的空间数据质量模型是采用层次化抽象方法建立的,以便逐渐趋近产品的技术特性与指标.

2.1 直接评价质量模型

从用户的角度出发,直接评价质量模型应包括可用性、可共享性和可加工性三个方面. 可用性就是指空间数据在精度、现势性、逻辑一致性、要素完备性等技术指标方面能否满足用户应用的要求;可共享性反映的是数据存储格式的通用性和开放性;可加工性则表示要素分类编码的合理性,即分类编码标准化的程度或反映现实世界真实性的程度. 直接评价质量模型由用户层和技术指标层两个层次来表达(如图3所示).

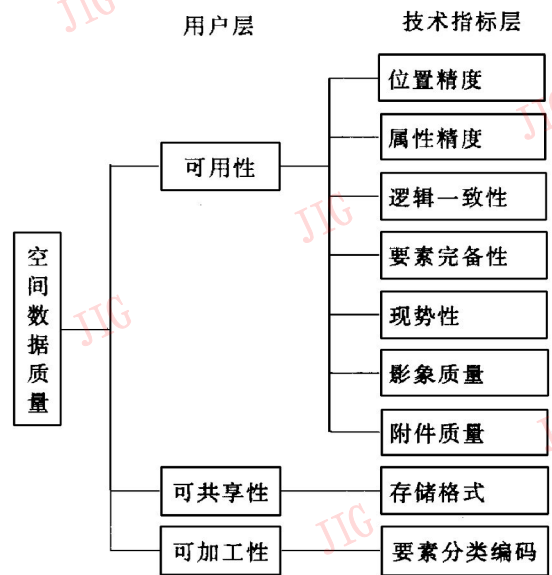


图3 直接评价质量模型

2.2 间接评价质量模型

间接评价的依据是一些间接信息,没有直接面对最终的数据集,其质量模型的选取只能以间接信息来确定,而这些间接信息也只能按照数据集形成的数据流中误差传播的途径来选择. 该质量模型通常有数据源质量、数据生产质量和数据加工处理质量三个层面. 其中数据源质量又可进一步细分为源图形(象)数据质量、源属性数据质量、文档数据质

量;而数据生产质量则包括数字化质量和数据转换质量;而数据生成、处理质量则分为计算误差、拓扑分析质量和图层叠置质量(如图 4 所示)。

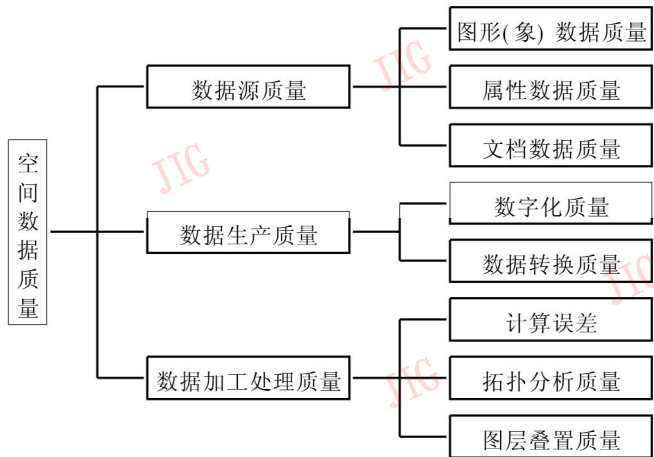


图 4 间接评价质量模型

2.3 直接评价质量模型与间接评价质量模型的统一

由图 3、图 4 可以看出,两个质量模型的最底层不是一个统一的质量元素层,为了形成最终统一的质量度量模型,以便为用户提供一致性的、非多义性的质量报告,必须使两个模型统一起来.但直接质量评价模型中表现用户的可用性和可加工性的技术指标,对于数据集来说,已反映在数据集的元数据中,无须评价,当用户需要时,可直接从元数据库中查询这些信息.因此,通过间接评价质量模型二级质量元素层的分析,其各项元素与直接评价质量模型的技术指标层存在着紧密的关联关系,如直接评价模型中的位置精度即是由间接评价质量模型中的源图形(象)数据质量、数字化质量、数据转换质量、计算误差、拓扑分析质量、图层叠置质量的积累.据此形成图 5 所示两个模型间的连接关系。

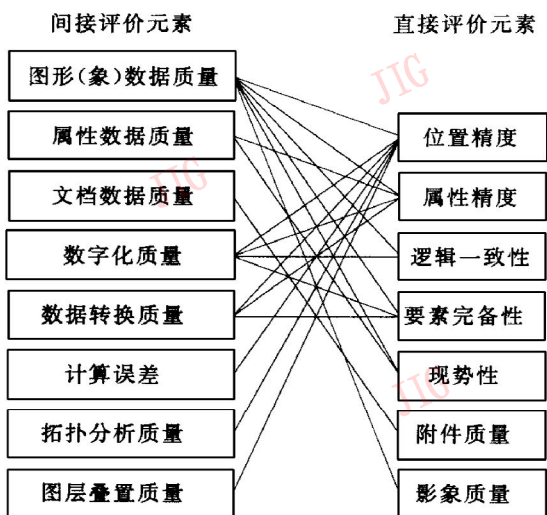


图 5 直接评价质量模型与间接评价质量模型的关系

3 空间数据质量度量模型

根据直接评价质量模型与间接评价质量模型间的关系,将直接评价质量模型与间接评价质量模型统一起来后,就可用其中一个子列的质量元素来统一度量数据质量,也就是说,不论采用什么质量评价方法都可以用一致的元素来度量空间数据质量.由于直接评价方法使用得更多,质量表达得更为准确,更适于用户判定产品适用性,因此我们选用直接评价元素来作为数据质量度量的元素,并建立如图 6 所示的空间数据质量度量模型。

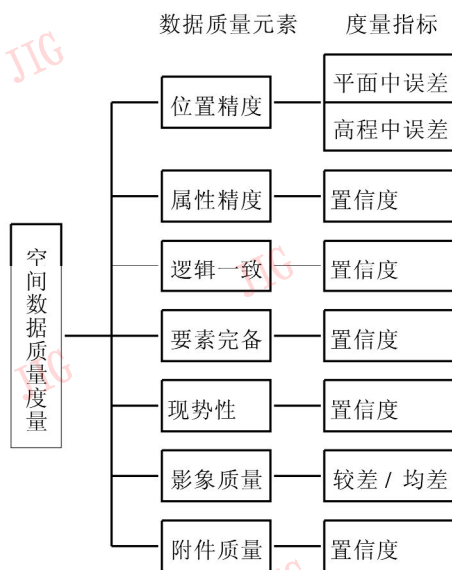


图 6 空间数据质量度量模型

4 结束语

本文提出的空间数据质量模型、质量度量模型、质量评价模型均属于概念级模型,它们均用于空间数据质量标准的制定,其中质量度量指标(如置信度、误差大小)在有关标准中规定.该模型中的质量元素的不同组合,可用于不同类型空间数据的质量评价,例如,矢量数据质量评价可采用位置精度、属性精度、逻辑一致性、要素完备性、现势性和附件质量作为质量元素;而影象数据质量评价可采用位置精度、影象质量和附件质量作为质量元素;格网数据质量评价则可采用属性精度和附件质量作为质量元素.如 1: 1 万、1: 5 万基础地理信息数据属于矢量数据,因而在对其采集与更新工作中我们使用该模型对其进行了一定程度的试验,从结果看基本反映

了空间数据质量的内在特征. 由于空间数据质量影响因素的多样性, 一些度量指标的非定量性, 因此对空间数据质量的度量指标的大小与评价方法还必须进一步深入研究.

参 考 文 献

- 1 蒋景瞳、何建邦、赵艳华等. 城市地理信息系统标准化指南. 北京: 科学出版社, 1998, 158~ 168.
- 2 刘文宝等. 平面线位误差带几何形状的解析表达式. 测绘学报, 1998, 27(3): 231~ 237.
- 3 蒋景瞳等. 美国国家数字制图数据标准. 北京: 测绘出版社, 1990. 11.
- 4 Dai Honglei, Liu Wenbao, Du Daosheng. A united model of visualizing positional uncertainties for spatial objects within vector GIS, In: Proceedings of The International Symposium on Spatial Data Quality, Hongkong: The Hong Kong Polytechnic University. 1999, 278~ 284.
- 5 Tong Xiaohua, Shi Wenzhong, Liu Dajie, Developing of error model for transition curves in GIS, In: Proceedings of The International Symposium on Spatial Data Quality, Hongkong: The Hongkong Polytechnic University, 1999, 23~ 32.
- 6 Goodchild M F, Hunter G J. A simple positioning accuracy measure for linear features, International Journal of Geographic Information science, 1997, (11): 299~ 306.
- 7 杜道生. GIS中空间数据质量与空间数据标准, RS·GPS·GIS的集成与应用. 北京: 测绘出版社, 1995, 63~ 73.
- 8 史文中. 地理信息系统中几何特征不确定性的通用模型: 从一维到N维. 测绘学报, 1997, 26(2): 160~ 167.
- 9 Guptill S C, Morrison J L. Elements of spatial data quality, Oxford: Persimmon, 1995.
- 10 Aalders H J G L. The registration of quality in GIS, In: Proceedings of The International Symposium on Spatial Data Quality, Hongkong: The Hong Kong Polytechnic University. 1999, 23~ 32.



杜道生 教授, 博士生导师, 57岁, 现任武汉测绘科技大学测绘遥感信息工程实验室空间信息系统研究室主任, 主要从事空间数据质量和地理信息标准化研究. 承担和参与多项国家和省部级科研项目, 获奖多项, 发表论文50余篇.

王占宏 高级工程师, 博士, 32岁. 现任国家测绘局测绘标准化研究所所长助理. 主要从事地理信息系统和地理信息标准化研究, 已负责制定过多项数据质量标准.

马聪丽 工程师, 28岁. 在国家测绘局测绘标准化研究所从事地理信息标准的研究工作.

敬告订户

《中国图象图形学报》从今年起已正式改为半月刊, 其中下半月出版的学报为全彩色版. 并由本刊编辑部自己发行, 现已开始证订, 全年订费与月刊相同, 而且凡订上半月学报者, 赠送下半月学报, 订阅下半月学报者, 赠阅上半月学报. 由于部分在邮局订阅的读者, 本刊不掌握通信地址, 因此为了使邮局订户能收到全年的刊物, 麻烦您将通信地址(邮编)及订刊复印件寄到本编辑部, 以便给您及时寄刊. 同时也希望您能给学报多提些建议并提出您的要求.

通信地址: 北京市海淀区花园路6号(100088)

《中国图象图形学报》编辑部 发行部收